VIROLOGY 212, 734–740 (1995)

SHORT COMMUNICATION

Sequence Analysis of the UL39, UL38, and UL37 Homologues of Bovine Herpesvirus 1 and Expression Studies of UL40 and UL39, the Subunits of Ribonucleotide Reductase

CLAIRE SIMARD,*,[1] ISABELLE LANGLOIS,* DOMINIK STYGER,† BERND VOGT,† CESTMIR VLCEK,‡ ANICK CHALIFOUR,* MICHEL TRUDEL,* and MARTIN SCHWYZER†

*Centre de Recherche en Virologie, Institut Armand-Frappier, Université du Québec, 531 Boulevard des Prairies, Laval des Rapides, Ville de Laval, Québec, Canada, H7V 1B7; †Institut für Virologie, Veterinärmedizinische Fakultät, Universität Zürich, Winterthurerstrasse 266a, CH-8057 Zürich, Switzerland; and ‡Institute of Molecular Genetics, Czech Academy of Sciences, Flemingovo Nam. 2, CZ-16637 Prague 6, Czech Republic

Received May 5, 1995; accepted July 13, 1995

We sequenced the region of the bovine herpesvirus type 1 (BHV-1) genome corresponding to map units 0.172–0.230 (7964 bp), representing the UL39, UL38, and UL37 homologues of herpes simplex virus which encode the large subunit of ribonucleotide reductase (RR) and components of the viral capsid and the tegument, respectively. To discriminate between two potential initiator AUGs of the UL39 gene, the 5′ end of the mRNA was mapped by S1 nuclease protection assays. Comparison of the amino acid sequences of the three BHV-1 proteins with analogous polypeptides from several other herpesviruses revealed significant levels of homology. We also compared the expression kinetics of the large (R1, UL39) versus the small (R2, UL40) RR subunits during the course of *in vitro* BHV-1 infection by Western blotting using specifically developed and calibrated antisera. Our results show that the R1 protein was synthesized earlier than its R2 counterpart. Moreover, the R1 protein accumulated to a higher level than the R2 protein even though the R2 transcript was in greater abundance than the R1 mRNA. This is discussed with regard to the translational efficiency of their transcripts.

Bovine herpesvirus 1 (BHV-1), a member of the *Alphaherpesvirinae,* is an economically important pathogen of cattle. BHV-1 is primarily associated with two major clinical syndromes of cattle, namely infectious bovine rhinotracheitis (IBR) and infectious pustular vulvovaginitis. The viral genome is a 137-kb linear double-stranded DNA which is believed to encode at least 70 polypeptides from which 25–33 are structural components of the virion. It also encodes transcriptional regulatory proteins and enzymes involved in DNA metabolism (*1*), such as ribonucleotide reductase (RR). This enzyme catalyzes the reduction of ribonucleoside diphosphates to their corresponding 2′-deoxy forms which are precursors for DNA synthesis. In herpes simplex virus type 1 (HSV-1), the holoenzyme is constituted of two large (R1) and two small (R2) subunits (*2*).

Presently, an international collaboration is under way to sequence the complete BHV-1 genome. As a contribution to this collaborative effort, our laboratories have sequenced the region from 0.172 to 0.230 map units (mu) from IBR-like isolates 34 (*3*) or Cooper (*4*). Together with the previously sequenced 0.162–0.172 mu region (*5*), it constitutes a 9.4-kb *Kpn*I–*Eco*RI fragment spanning the *Hin*dIII fragments I and E of the genome (Fig. 1). Computer analysis revealed that this 9.4-kb fragment encoded the BHV-1 homologues of the HSV-1 UL40, UL39, UL38, and UL37 gene products, with the first three genes in leftward orientation and the last one running toward the right. The four coding sequences are thus directed in the same orientation as in HSV-1 (*6*), considering that the HSV-1 prototype genome is conventionally inverted relative to the one from BHV-1 (*7*). A fifth ORF of significant length was found inside the UL39 ORF with a frameshift of +2, potentially encoding 317 amino acid residues. However, codon usage analysis and a comparison with the corresponding regions of other alphaherpesviruses, all of which were found to carry stop codons in the +2 frame, made it unlikely that this frame encodes a protein. The 7964-bp sequence presented in Fig. 2 begins nine codons from the N-terminus of the previously reported sequence of the gene encoding the small subunit (R2) of the BHV-1 RR (*5*).

An ORF initiating at position 2721 and ending at base 49 was identified and its corresponding amino acid sequence was found to be homologous to the R1 of other viral RR. This ORF would encode a polypeptide with a calculated molecular weight of 96,750 Da. However, translation could initiate at a second Met codon (position 2409), still producing a protein large enough (86,004 Da)

[1] To whom correspondence and reprint requests should be addressed. Fax: (514) 686-5626.

0042-6822/95 $12.00

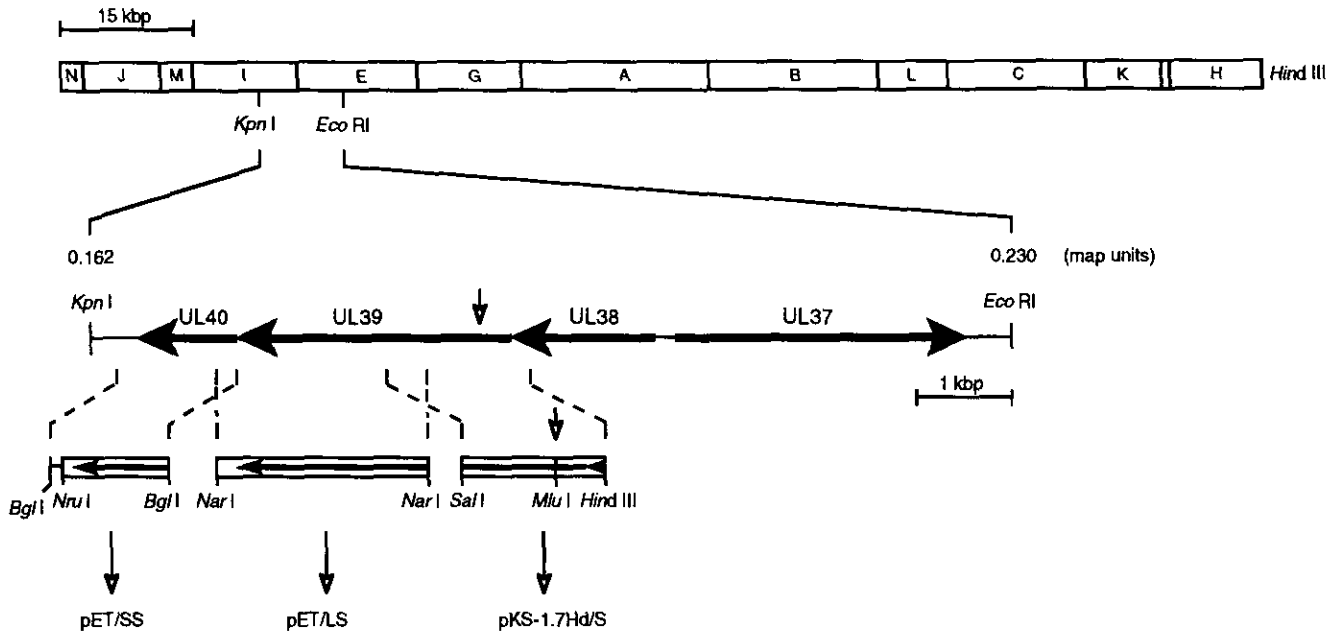

FIG. 1. Location and orientation of the UL40–UL37 homologues on the BHV-1 genome. The *Hin*dIII restriction map of the 137-kb genome of BHV-1 (*4*) is illustrated at the top. The 9.4-kb *Kpn*I/*Eco*RI fragment corresponding to map units 0.162–0.230 is shown in an expanded form. Arrows indicate the location and orientation of open reading frames homologous to UL40, UL39, UL38, and UL37 HSV-1 genes. The complete DNA and amino acid sequences of the three latter genes are given in Fig. 2; those for the UL40 were reported previously (*5*). The small arrow above the UL39 ORF identifies the translation start site of the corresponding polypeptide (see text). The R2 (UL40) and R1 (UL39) proteins of viral RR were individually expressed in *Escherichia coli* using the pET21b translation vector (Novagen, Inc.) to serve as antigens for the production of monospecific sera. To express the R2 polypeptide, the 1180-bp *Bgl*I fragment representing its complete coding sequence with the exception of the first 2 amino acid residues was ligated to a *Bam*HI adaptor, digested with *Bam*HI and *Nru*I, then inserted in frame with the T7-Tag coding region of pET21b to generate pET/SS. To express the R1 protein, the 1992-bp *Nar*I fragment encoding its last 602 amino acid residues was ligated to a *Bam*HI adaptor, digested with *Bam*HI, and inserted in frame with the T7-Tag region of pET21b, generating pET/LS. In order to ascertain proper reading frames, the vector/insert junctions of both pET/SS and pET/LS clones were sequenced. The subclone pKS-1.7 Hd/S was used to map the 5′ end of the UL39 transcript. Diagrams are drawn to scale as indicated.

to represent the R1 of RR. To determine at which AUG translation initiates, the 5′ end of the R1-specific transcript was mapped by S1 nuclease protection assays (Fig. 3). For this purpose, the 760-bp *Mlu*I–*Hin*dIII fragment which was end-labeled 15 nucleotides downstream from the first base of the AUG_{2409} (see Fig. 2) was hybridized with total RNA extracted either from BHV-1-infected (9 hr postinfection) or mock-infected MDBK cells. Following S1 nuclease digestion, three DNA fragments of 222, 223, and 224 bases long were protected in the reaction mixture containing BHV-1 mRNA (Fig. 3, lane BHV), whereas no fragments were protected in the control (lane m-i). The S1-resistant DNA band of 223 bases was twice as intense as the other two, suggesting that it corresponds to the major transcription start site which is located 209 bases upstream from the translation initiator codon AUG_{2409} (position 2618 in Fig. 2). The other two bands may either represent partially digested DNA fragments or minor transcription start sites at 208 and 210 nucleotides from the AUG_{2409}. These findings excluded the possibility that the AUG_{2721} was used as the initiator codon since transcription was initiated downstream from it. We previously reported that the transcript encoding the R1 is 3′-coterminal with the one encoding the

FIG. 2. Nucleic acid and deduced amino acid sequences of the UL39, UL38, and UL37 homologues of BHV-1. DNA sequencing was performed either by the dideoxynucleotide chain termination (*18*) or by the Maxam and Gilbert (*19*) method. The DNA sequence is shown as the rightward 5′ to 3′ strand and its nucleotides are numbered on the right. Encoded protein sequences are shown below the DNA sequence and the corresponding amino acid residues are numbered in bold. The direction of translation of the corresponding genes is identified with arrows. The UL39 ORF extends from nucleotide 2721 to 49; however, the actual translation start site of the polypeptide is at position 2409. The transcriptional start site of UL39 and its direction are indicated with a vertical bar and an arrow at position 2618. The coding sequences of the UL38 and UL37 homologues are located within the regions 4175–2754 and 4391–7462, respectively. Putative TATA and CAAT boxes upstream from each coding sequence are underlined with a double line. The putative polyadenylation signal of UL38 is identified with a single line above the nucleotide residues. Restriction sites which are pertinent to this study are underlined and identified. Every 10th base in a line is denoted with a dot over the corresponding residue. The nucleotide sequence data have been submitted to the EMBL nucleotide sequence database and have been assigned Accession No. Z49078. The sequence will become available to the scientific community during the course of 1995 when the sequencing of the BHV-1 genome will be completed.

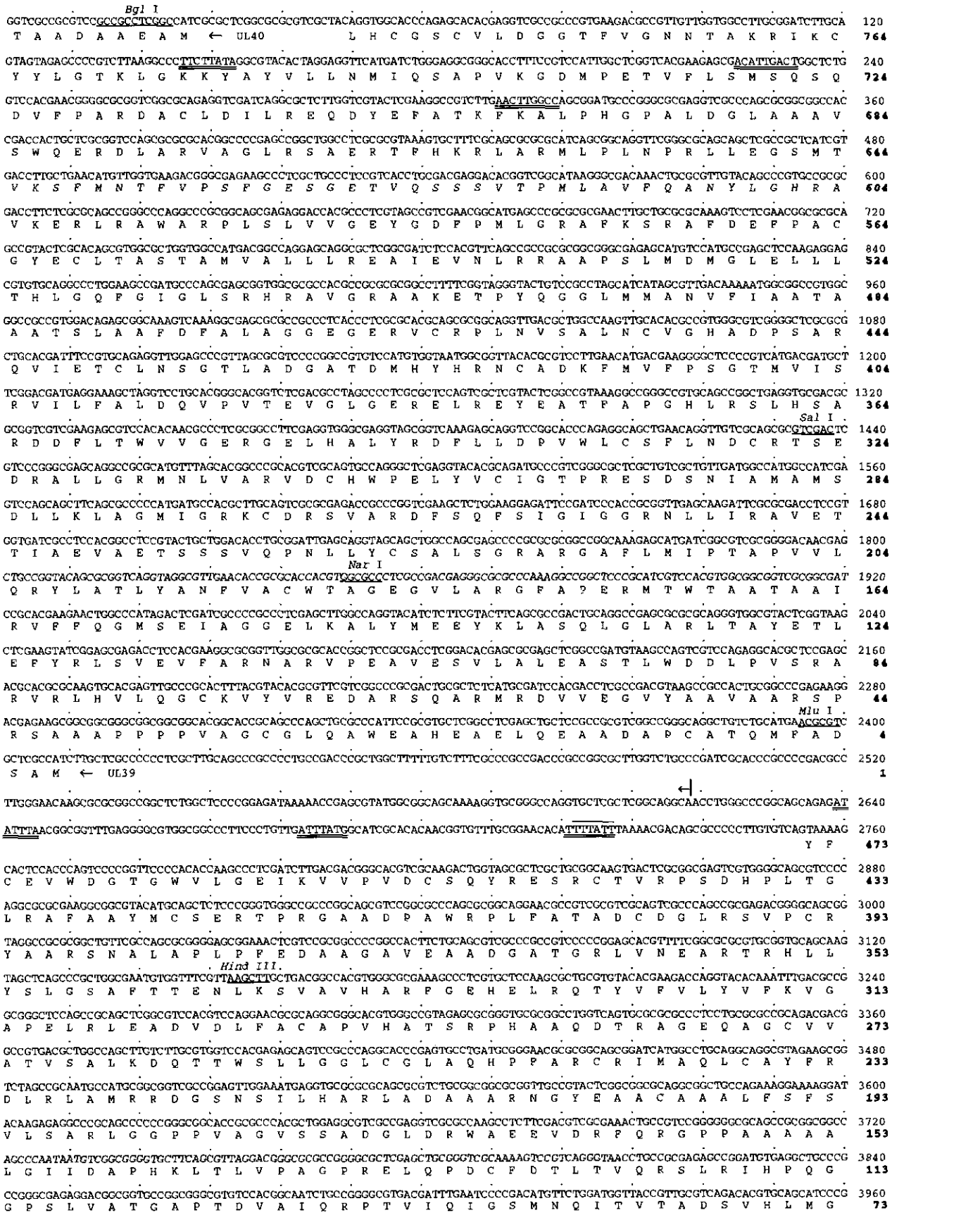

FIG. 2

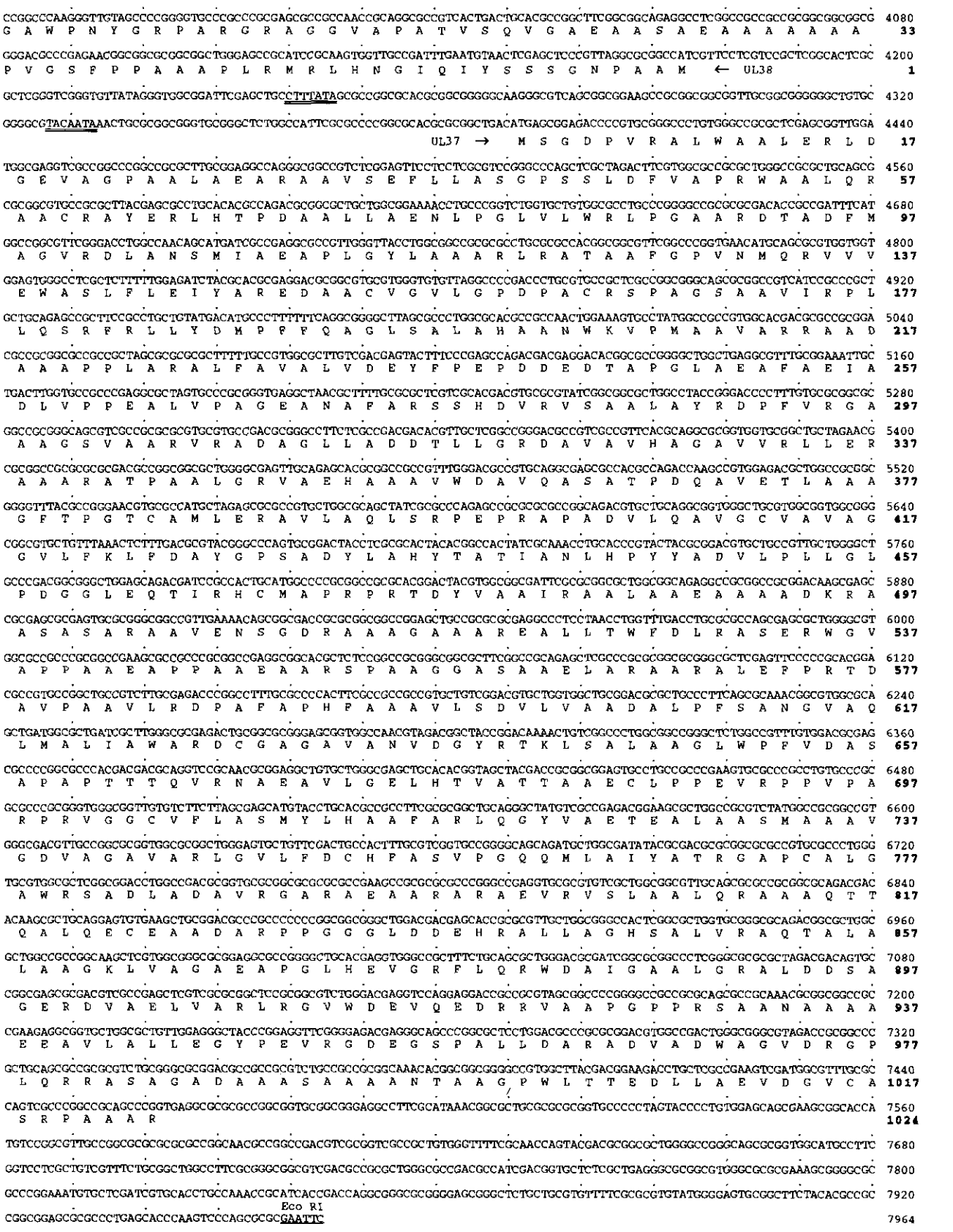

FIG. 2—*Continued*

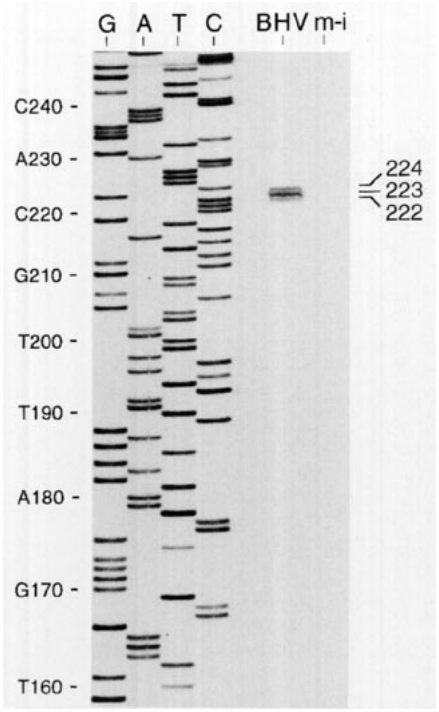

FIG. 3. 5′ end mapping of the transcript encoding the large subunit of BHV-1 RR. The S1 nuclease protection assays were performed as previously described (*5*). The 760-bp *MluI*–*Hin*dIII fragment end-labeled at the *MluI* site (position 2394, see Fig. 2) was isolated from pKS-1.7Hd/S (Fig. 1) and hybridized to RNAs isolated as previously described (*20*) from BHV-1- or mock-infected cells. DNA from the BHV-1 (lane BHV) or mock-infected (lane m-i) mixture which was resistant to S1 nuclease digestion was analyzed on sequencing gels in parallel with dideoxy sequencing reactions of M13mp18 (lanes G, A, T, C). Every 10th nucleotide of the M13mp18 sequence is identified on the left with the size of the corresponding fragment. The sizes of S1-resistant DNAs in the BHV-1 sample are identified on the right.

enzyme's R2 and that its 3′ end site is located 18 bases downstream from a shared polyadenylation signal (*5*). This places the site 1050 bases downstream from the stop codon of the gene. Adding the sizes of the 5′ and 3′ noncoding regions of the transcript to that of its coding sequence would result in a transcript of 3623 nucleotides long. This correlates well with the observed transcript's size (3.7 kb) following Northern blot analysis (*5*).

The R1 transcript is presumably regulated by a TATA box located 21 nucleotides upstream from the transcription start site. Two additional TATA boxes located at 65 and 103 nucleotides and a CAAT box motif 1100 bases upstream are probably too distant to serve as promoter elements. By comparison, the putative promoter of the transcript encoding R2 is composed of a single TATA and two CAAT boxes (Fig. 2). Interestingly, we previously reported that the R1 transcript is at least five times less abundant than the one encoding R2 (*5*). Although this may reflect differences in the transcripts' turnover, differences in the promoter elements of the two genes could account for the overexpression of the R2 transcript relative to the other one.

Significant homologies were found between the R1 protein of BHV-1 and those of the three herpesvirus subfamilies (*Alpha-, Beta-,* and *gammaherpesvirinae*), of an iridovirus, and of mammalian origin, supporting the hypothesis that the genes were derived from a common precursor. Homologies were distributed throughout the proteins with the exception of the N-terminal ends where the similarities were weak. We identified several conserved domains which may reflect structural constraints indispensable to the enzyme's active sites or to the assembly of holoenzymes. Consistent with this hypothesis, residues which have been implicated in *Escherichia coli* R1 active sites or in ribose binding and transformation (Cys225, Cys462, Cys439, Tyr730, Tyr731, Ser224, Asn437, and Glu441; *8*) were strictly conserved in all sequences examined except in those from human cytomegalovirus and human herpesvirus type 6.

Examination of the BHV-1 nucleotide sequence to the right of the RR genes revealed the presence of the UL38 and UL37 ORFs in the same divergent arrangement as that of other alphaherpesviruses. The UL38 ORF extends from AUG_{4175}, which is in a good sequence context for translation initiation (*9*), to UGA_{2753} and encodes a polypeptide of 474 amino acids. It is flanked by a TATA box (positions 4246–4240) and a polyadenylation signal (positions 2727–2722). According to the location of these elements, we predicted a UL38 transcript of 1.7 kb including the poly(A) tail. An early transcript of this size has indeed been observed by Northern blot analysis of BHV-1 strain K22 (*10*). Furthermore, it precisely maps in the expected position because it was shown to span *Hin*dIII fragments E and O; the 1.4-kb fragment O is present only in strain K22 due to an additional *Hin*dIII site at 0.184 mu and corresponds to the right end of *Hin*dIII fragment I in the Cooper strain (*10*).

The UL37 ORF comprises positions 4391–7462 (1024 amino acid residues). Potential TATA boxes located upstream of UL37 deviate from the consensus at several positions. There is no downstream polyadenylation signal within the sequence presented here, and the adjacent UL36 sequence is not yet available. Nevertheless, a 3.9-kb transcript and a very long transcript (at least 8 kb) have been mapped by Northern blot analysis to the *Hin*dIII E region (*10*) and may encode UL37 and UL36, respectively, although precise transcript boundaries have not been established.

Multiple sequence alignments of the BHV-1 UL38 and UL37 ORFs revealed strong homologies to the corresponding proteins of all alphaherpesviruses which have been sequenced and weaker homologies to the proteins of beta- and gammaherpesviruses. In HSV-1, both the UL38 and UL37 gene products are associated with virions and have been shown to be essential for virus replication in cell culture (*11*). The UL38 protein participates in icosahedral capsid assembly (*12, 13*); the UL37 protein is associated with the viral tegument (*14, 15*) and also

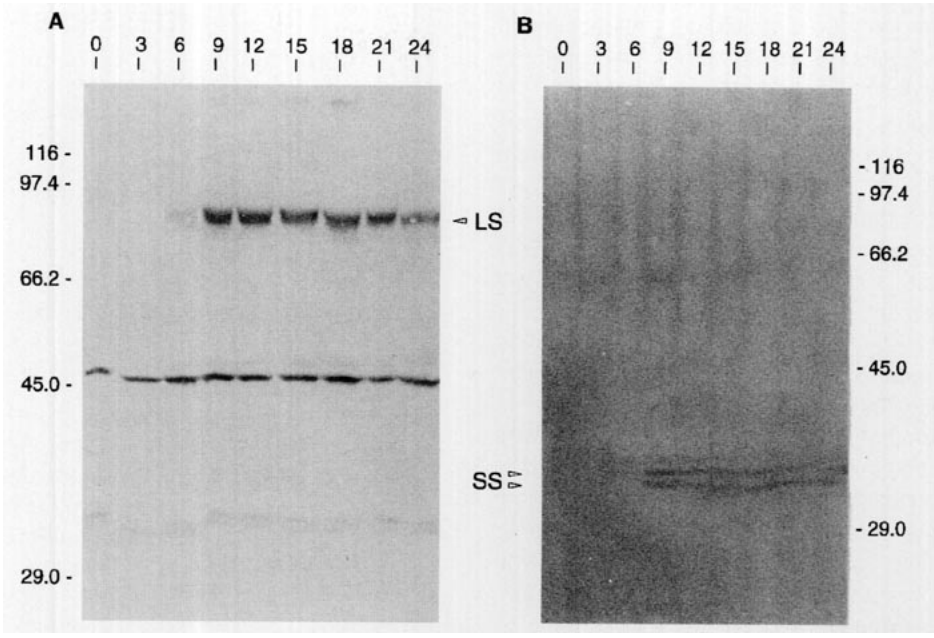

FIG. 4. Expression kinetics of the large and small subunits of BHV-1 RR during the course of *in vitro* infection. Confluent monolayers of MDBK cells (18×10^6 total cells) were infected with BHV-1 at an m.o.i. of 2 for 0, 3, 6, 9, 12, 15, 18, 21, and 24 hr. Cells were washed with PBS and lysed in 500 μl of SDS sample buffer (40 m*M* Tris–HCl, pH 6.8, 2% SDS, 280 m*M* β-mercaptoethanol, 10% glycerol). Lysates (40 μl) were fractionated by SDS–PAGE (10%) and proteins were electrotransferred at 55 V for 4 hr onto a PVDF membrane (Bio-Rad), conditions which allowed complete transfer of high MW proteins. Western blots were performed as described (*21*) using antisera specific to either (A) the R1 or (B) the R2 of BHV-1 RR. Antisera against R1 and R2 were obtained from guinea pigs immunized with the respective polypeptides purified on SDS–PAGE from extracts of *E. coli* BL21 harboring the recombinant plasmids pET/SS and pET/LS (Fig. 1), respectively. Reacting bands were revealed with protein A peroxidase. LS and SS specify the large (R1) and small (R2) subunits of BHV-1 RR, respectively. Sizes in kDa of molecular weight standards are identified on the left and right.

interacts with ICP8, the major DNA binding protein (*16*). Interestingly, the HSV-1 UL38 gene is regulated with strict late kinetics (*17*), whereas the BHV-1 UL38 transcript belongs to the early kinetic class (*10*).

The RR coding sequences obtained in this and a previous study (*5*) allowed us to express the individual subunits as T7-Tag fusion proteins in *E. coli*, via pET/LS and pET/SS recombinant translation vectors (Fig. 1), for the generation of specific antisera. These sera were used to verify whether the accumulation of the two polypeptides in BHV-1-infected cells correlated with our previous observation that the R2 transcript is more abundant than that of R1 (*5*). The small (pET/SS) and large (pET/LS) subunit-specific vectors allowed the synthesis of abundant polypeptides of 36 and 67 kDa, respectively (results not shown), which were of the expected sizes (36,965 and 67,600 Da, respectively). The two recombinant proteins reacted specifically with the commercially available T7-Tag monoclonal antibody (MAb), demonstrating that they included the T7-Tag N-terminal region provided by the translation vector (results not shown). Both proteins were purified by SDS–PAGE and used to generate antisera. Since differences in the immunogenicity of the antigens would interfere with the evaluation of the relative abundance of the two subunits of viral RR in infected cells, the relative concentrations of the antisera were evaluated. For this purpose, we took advantage of the common T7-Tag region of the two fusion polypeptides expressed in *E. coli*. Insoluble fractions obtained from *E. coli* lysates containing either the R2 or R1 hybrid proteins were analyzed by Western as well as dot blottings (results not shown). Since the T7-Tag MAb reacted with the two fusion proteins, it was used as an internal standard to calibrate their relative abundance when using their specific antisera. Dilutions of the antisera specific to the R2 and R1 were adjusted to obtain a difference in the signal which was empirically equivalent to that obtained with the T7-Tag MAb.

Accumulation of the R2 and R1 of BHV-1 RR was analyzed as a function of time by Western blotting of infected cell lysates (Fig. 4). The R1-specific antiserum reacted with a viral polypeptide of 85 kDa in BHV-1-infected cell lysates (Fig. 4A, lanes 3–24) which was absent from mock-infected cells (lane 0). The size of the polypeptide is consistent with the one predicted (86,004 Da) from the deduced amino acid sequence of R1, starting at the AUG_{2409}. The R1 polypeptide was detectable, albeit at a low level, as soon as 6 hr postinfection, indicating that it is expressed as an early product (β) following cell infection. Its levels significantly increased up to 9 hr and remained steady until 24 hr postinfection. The three other proteins (45, 35, and 33 kDa) which were revealed in BHV-1-infected as well as in mock-infected lysates represent cellular proteins cross-reacting with the antiserum.

The R2 antiserum specifically reacted with two polypeptides of 34 and 33 kDa in BHV-1-infected cell lysates (Fig.

4B, lanes 3–24) but did not recognize any polypeptides in mock-infected cells (lane 0). The size range of the polypeptides observed is consistent with the one predicted from the amino acid sequence of the viral R2 (35.25 kDa; *5*). Although the lower band observed may represent a degradation product of the larger one, the possibility that it may represent a different form of the R2 cannot be excluded. The polypeptides accumulated in cells at similar levels as of 9 hr postinfection and their levels remained constant until 24 hr postinfection.

In contrast to the transcript levels, the R1 polypeptide was more abundant than that of the R2. Furthermore, even though at 6 hr postinfection the R2 transcript is five times more abundant than that of R1 (*5*), the R2 polypeptide was only detected at 9 hr, whereas the R1 protein was present as early as 6 hr. Together, these results strongly suggest that the transcript encoding R1 is more efficiently translated than that of R2. Based on Kozak's findings (*9*), the presence of an A residue at position −3 of the initiator AUG codon of the R1 transcript is preferable for translation to the G residue found at this position in R2. Additionally, the larger size of the R1 transcript leader sequence (209 bases relative to 95 for the R2; *5*) may also be advantageous. These features present in the R1 transcript could facilitate its translation and explain the abundance of R1. However, the difference in the abundance of the R1 polypeptide compared to the R2 could also be explained by differences in their half-lives. Since in every organism in which ribonucleotide reductase has been characterized the stoichiometry of the holoenzyme's two subunits is 1:1, our results strongly suggest that free R1 subunits or R1 homodimers exist in the cell.

ACKNOWLEDGMENTS

The authors thank Dr. Johnny Basso for proofreading the manuscript. This work was funded by NSERC OGP0006211 and OGP0003579 as well as by Swiss National Science Foundation Grant 92-03791.

REFERENCES

1. Wyler, R., Engels, M., and Schwyzer, M., *In* "Herpesvirus Diseases of Cattle, Horses, and Pigs" (G. Wittmann, Ed.), pp. 1–72. Kluwer Academic, Boston/Dordrecht/London, 1989.
2. Ingemarson, R., and Lankinen, H., *Virology* **156,** 417–422 (1987).
3. Simard, C., LaBoissière, S., Séguin, C., and Trudel, M., *Intervirology* **32,** 117–126 (1991).
4. Mayfield, J. E., Good, P. J., Vancort, V. H., Campbell, A. R., and Reed, D. E., *J. Virol.* **47,** 259–264 (1983).
5. Simard, C., Bastien, N., and Trudel, M., *Virology* **190,** 689–701 (1992).
6. McGeoch, D. J., Dalrymple, M. A., Davison, A. J., Dolan, A., Frame, M. C., McNab, D., Perry, L. J., Scott, J. E., and Taylor, P., *J. Gen. Virol.* **69,** 1531–1574 (1988).
7. Davison, A. J., and Wilkie, N. M., *J. Gen. Virol.* **64,** 1927–1942 (1983).
8. Uhlin, U., and Eklund, H., *Nature* **370,** 533–539 (1994).
9. Kozak, M., *J. Biol. Chem.* **266,** 19867–19870 (1991).
10. Wirth, U. V., Gunkel, K., Engels, M., and Schwyzer, M., *J. Virol.* **63,** 4882–4889 (1989).
11. Ward, P. L., and Roizman, B., *Trends Genet.* **10,** 267–274 (1994).
12. Tatman, J. D., Preston, V. G., Nicholson, P., Elliott, R. M., and Rixon, F. J., *J. Gen. Virol.* **75,** 1101–1113 (1994).
13. Thomsen, D. R., Roof, L. L., and Homa, F. L., *J. Virol.* **68,** 2442–2457 (1994).
14. McLauchlan, J., Liefkens, K., and Stow, N. D., *J. Gen. Virol.* **75,** 2047–2052 (1994).
15. Schmitz, J. B., Albright, A. G., Kinchington, P. R., and Jenkins, F. J., *Virology* **206,** 1055–1065 (1995).
16. Shelton, L. S., Albright, A. G., Ruyechan, W. T., and Jenkins, F. J., *J. Virol.* **68,** 521–525 (1994).
17. Guzowski, J. F., Singh, J., and Wagner, E. K., *J. Virol.* **68,** 7774–7789 (1994).
18. Sanger, F., Nicklen, S., and Coulson, A. R., *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467 (1977).
19. Maxam, A., and Gilbert, W., *Proc. Natl. Acad. Sci. USA* **74,** 560–564 (1977).
20. Simard, C., Nadon, F., Séguin, C., LaBoissière, S., and Trudel, M., *Arch. Virol.* **110,** 63–75 (1990).
21. Trudel, M., Nadon, F., Séguin, C., Boulay, G., and Lussier, G., *Vaccine* **5,** 239–243 (1987).